

GIS Data Specifications for Resource Mapping, Inventories, and Studies

FINAL DRAFT



National GIS and Inventory & Monitoring Programs
National Park Service

Summary

Resource Management (RM) and Inventory and Monitoring (I&M) projects and activities generate both spatial and tabular data sets. These GIS data sets will be incorporated into park, regional, and national databases and made available to a wide range of users. In order for this to occur effectively, certain standards and product specifications must be adhered to. This document provides general standards for spatial data collection and submission. Park-, network-, region-, and program-level project managers may require further specifications and must approve any deviation from these standards.

Deliverables

Complete and verified data will be delivered via CD-ROM (preferred) and/or by software compressed (zipped) file. All digital and hardcopy information that is part of the project must be delivered, including GIS data reports, metadata, photos, and other supporting materials. Each CD should be in CDR format, so that once it is written it cannot be modified. The CD should be in ISO 9660 format to allow cross-platform use (this requires 8.3 file names). The products delivered to the project manager will contain the following items:

Required

- Descriptive document
- Spatial data
- Associated data table(s) or relational MS Access Database
- FGDC-compliant Metadata

As Specified

- ArcView 3.x Legend (.AVL file)
- GIS Theme Manager Theme Lists
- Linked document(s)
- Linked graphics or digital photographs

Descriptive Document

A Microsoft Word document (and/or ASCII text file if specified) describing the data set will accompany any submission and provide all necessary information for understanding the submittal. This includes but is not limited to the following:

- Contents of the CD or .zip file
- Sensitive data issues (if any exist)
- Concise summary of accuracy assessment procedures applied
- Recommended "official" theme name (or filename alias)
- Contact information for those responsible for the data
- Data dictionary for all attribute and database tables (e.g., listed by table in "field name", "data type", "data width", "field description" tabular format)
- Linking fields (to documents, MS Access database, digital photographs)
- Viewing scale thresholds (if applicable)

The following example of a Descriptive Document for a park with alpha code "*CODE*" can be used as a template.

CODE_BirdSurvey.Doc

A CDR in ISO 9660 format contains the following file:

CODEBird.Zip containing the following files:

- *CODE_BirdSurvey.Doc* (this descriptive document)
- *CODE_Bird_File_Names.Doc* (naming convention or codes used for file names - if applicable)
- *CODE_BirdSurvey2000.Doc* – a descriptive document for the *Code* 2000 bird survey.
- *CODEBird.e00* – exported ArcINFO Coverage
- *CODEBird.avl* – ArcView 3x legend file
- *CODEBird.txt/html/.sgml* – FGDC metadata formats
- *CODEBird.mdb* – MS Access database
- *CODE_Bird_Data_Dict.Doc*

None of the data contained in this data set is considered sensitive.

Features were marked in the field on 1:63,360 paper maps and digitized using a tablet. Digitized spatial data were plotted and compared to the original maps. Digitized points fell within 0.1 inches of the original marked points.

An appropriate Theme name for this data should contain Bird Survey and the year (2000) like – “Bird Survey 2000”

The data were created by Joe Smith of the National Park Service, Some NPS Project, phone – (999) 999-9999.

The data dictionary for attribute and data tables are included in the file *CODE_Bird_Data_Dict.Doc*

The Key Field “LocationID” links the Access database and the coverage.

No viewing scale thresholds are required for this data.

Spatial Data

There are several ways in which spatial data can be represented in a GIS including points, lines, polygons, or rasters/images (or geodatabases that are not recommended for contracted data development or general distribution). Determining which method(s) is appropriate for your study involves consideration of scale and study goals. Prior to data collection, this issue should be addressed and resolved in the project study plan in consultation with the project or data manager. Additionally, network and park data management plans may dictate the appropriate format.

Naming Conventions

A clear and meaningful filename should be used that conveys the nature of the data and park unit it relates to. File names should adhere to 8.3 naming standards and not contain spaces or special characters. Field names should be 10 characters or less to conform to Dbase and ArcINFO field naming limitations.

Coordinate Systems

All spatial data collected or submitted for national, regional, or network NPS programs shall be geo-referenced and provided in a standard projection. Digital geospatial data should be referenced to two coordinate systems--the current standard system used by the individual park (generally UTM, NAD83) AND a regional-scale system (generally Geographic, NAD83). The steps used to get the data into the proper projection must be documented in the metadata. The project manager must specify and approve any deviation from these projection standards.

NPS-wide and Regional Data Standard (one copy) The standard projection for most NPS regions and national programs is geographic with the following parameters:

• Datum	North American Datum 1983
• Spheroid	GRS 1980
• Units	Decimal degrees

Park Unit Data Standard (one copy) In general, the standard projection for most park-level GIS layers is Universal Transverse Mercator with the following parameters:

• Projection	Universal Transverse Mercator
• Datum	North American Datum 1983
• Spheroid	GRS 1980
• False Easting	500,000
• False Northing	0
• Units	Meters

Park Unit Standards for Exceptions In addition to the systems noted above, several NPS units require additional specific standards for data delivery (e.g., Cabrillo and Craters of the Moon National Monuments and Sequoia and Kings Canyon National Parks). If a park crosses UTM zone boundaries, it is recommended that only one zone, or a different coordinate system, should be used. Parks in Hawaii and other Pacific islands will be in the datum and projection specified by each park. Because of their unique geographic location, the NPS Alaska Region also requires a specific datum and projection as noted below.

Alaska Region The standard projection for Alaska Region parks uses the following parameters:

• Projection	Alaska Albers Equal Area
• Datum	North American Datum 1927
• Spheroid	Clark 1866
• False Easting	0
• False Northing	0
• Central Meridian	-154 00 00
• 1st Standard Parallel	55 00 00
• 2nd Standard Parallel	65 00 00
• Units	Meters

Data Formats

All vector data will be supplied as an ArcINFO coverage, ArcINFO interchange file (*.E00) and/or ArcView Shape file, compatible with the current version of ArcINFO. All raster data will be supplied as an ArcINFO GRID and ArcINFO interchange file, compatible with the current version of ArcINFO. In general, all digital imagery, such as scanned aerial photographs, is to be supplied as tagged image file format (.TIFF) files with the proper header file (or world file) for geo-referencing purposes. If special circumstances exist (such as large image files), other spatial data formats may be acceptable. If not specified directly in the contract or project proposal, the data format(s) should be clearly stipulated and agreed upon with contractors or cooperators before data collection and processing start. If there are questions about choosing a data format, converting between formats, or non-standard formats, contact the park, network, region, or program GIS/data managers.

After the data are produced, any specified ArcView legends and/or NPS GIS Theme Manager theme list files (see <http://www.nature.nps.gov/im/apps/thmmgr/index.htm>) should be developed and included with the project deliverables. Brief reviews of current spatial data formats are listed below.

ESRI ArcINFO Coverage Data developed in ArcINFO coverage format should be exported to an .E00 file (ArcGIS 8.x .E00 files should include the metadata .XML file from ArcCatalog). All coverages should be created as double precision data sets. If the data set was originally obtained in single precision, convert it to double precision before submitting. Well-defined projection properties of the coverage are mandatory.

ESRI ArcView Shape File Shape file format shall be used only when an ArcINFO coverage does not exist. The shape file format includes at a minimum the .SHP, .DBF, and .SHX files (ArcGIS .SHP files should include the metadata .XML file from ArcCatalog). A .PRJ (projection definition) file is required unless specified otherwise in the contract or project proposal.

AutoCAD file with world file or .DXF export Generally, this format is NOT recommended and will be accepted only upon the direct specification, approval, and documentation of the project and data managers. If used, an Autocad release 14 (release 2000 for ArcGIS 8.x) or lower drawing (.DWG) file with a georeferenced world file (.WLD) and/or AutoCAD release 13 .DXF format file with georeferenced coordinates should be required. Note that some Autocad features such as X references, splines, etc. may not be available in ESRI products.

ESRI GRID File This is the preferred format for raster data and particularly useful for images that contain attributes other than cell values. Generally, GRID themes should be delivered as .E00 files as stipulated above. However, for large raster data sets, ESRI recommends sharing GRID files as separate workspaces because .E00 files may be extremely large and unwieldy.

GeoTiff v1.0 A raster format with geo-referencing stored in the header of the file.

Tiff with world file Tiff files shall be geo-referenced and include the world file (.TFW).

ERDAS Imagine file Imagine files shall be geo-referenced. Pyramid files (.RRD) shall be included if available.

Other possible raster file formats that may be utilized natively as an ArcView theme include .BMP, .BSQ, .BIL, .BIP, ERMapper, IMPELL Bitmaps, Image Catalogs, .JPEG, MrSID, and Sun

Rasterfiles, but applicable header or world files must be used (which makes .BMP, .JPEG, and Sun Rasterfiles unacceptable). Again, the appropriate project manager(s) must approve any deviation from the preferred standards listed in bold above.

Collection methods

Several approaches to capturing digital data can be employed including digitizing features from maps or aerial photographs, and GPS (Global Positioning System) collection. The appropriate method should be determined in the study plan and in consultation with the project, resource, or data manager. Criteria for acceptable GPS coordinates should be based on the National Park Service GPS Data Collection Guidelines at <http://www.nature.nps.gov/im/gis/standards.htm>.

When digitizing features from maps or photographs, the source, scale, date, and methods (i.e., process steps) shall be recorded in the Metadata and discussed in the Descriptive Document. When using GPS collection, the GPS unit type, averaging method, post processing and other criteria shall be recorded in the Metadata and discussed in the Descriptive document.

Scale and Spatial Resolution

Vector Data New data should be compiled with an accuracy level better than U.S. National Map Accuracy Standards for a 1:24,000 product unless other requirements exist (e.g., larger, more-detailed or smaller, regional-scale data). Project planners should contact appropriate GIS or data management staff for specific scale and spatial resolution requirements for vector data, which should be clearly specified in the contract or cooperative agreement.

Digital Image Data and Aerial Photography Specific scale and spatial resolution requirements for image data should be specified in the contract or cooperative agreement, or the contractor should contact the project manager for clarification. For vegetation classification under the NPS/USGS vegetation classification and mapping program, the current recommendation is 1:12,000 color infrared aerial photographs with 60% overlap (endlap) and 30% sidelap.

Horizontal and Vertical Accuracy

All spatial data collected shall be analyzed for their spatial accuracy and shall meet or exceed the National Map Accuracy Standards for the particular scale intended (for more information see <http://mapping.usgs.gov/standards/>). Longitude and Latitude coordinates for geographic data should be recorded to a minimum 5 significant digits to the right of the decimal point and stored in double precision attribute or database fields. Any calculations done with location data should be done at double precision with the results rounded or truncated to the appropriate propagated error limits. All calculations and processing completed on the spatial data shall be reported in the metadata.

For maps on publication scales larger than 1:20,000, not more than 10 percent of the points tested shall be in error by more than 1/30 inch, measured on the publication scale; for maps on publication scales of 1:20,000 or smaller, 1/50 inch. These limits of accuracy shall apply to positions of well-defined points only. Well-defined points are those that are easily visible or recoverable on the ground: monuments or markers, such as benchmarks and property boundary monuments; intersections of roads and railroads; and corners of large buildings or structures (or center points of small buildings). In general, what is well defined will also be determined by what is plot-able on the scale of the map within 1/100 inch. Thus, while the intersection of two roads or property lines meeting at right angles

would come within a sensible interpretation, identification of the intersection of such lines meeting at an acute angle would not be practicable within 1/100 inch...

Vertical accuracy, as applied to contour maps on all publication scales, shall be such that not more than 10 percent of the elevations tested shall be in error by more than one-half the contour interval. In checking elevations taken from the map, the apparent vertical error may be decreased by assuming a horizontal displacement within the permissible horizontal error for a map of that scale. (USGS Fact Sheet 078-96, 1997)

The following table provides the allowable horizontal accuracy for some common scales:

Scale	Allowable Error	Scale	Allowable Error
• 1:40,000	33.8 meters (111 feet)	• 1:9,600	4.9 meters (16 feet)
• 1:31,680	16.1 meters (53 feet)	• 1:4,800	2.4 meters (8 feet)
• 1:24,000	12.2 meters (40 feet)	• 1:2,400	1.2 meters (4 feet)
• 1:20,000	10.1 meters (33 feet)	• 1:1,200	0.6 meters (2 feet)
• 1:12,000	6.1 meters (20 feet)		

Attribute Data

By their nature resource inventories and studies will generate complex data sets. All fields within the database supporting GIS layers should have names of 10 characters or less due to ArcView and Dbase limitations. Because the ArcINFO Coverage/Shape file format is not ideal for storage and management of complex relational data, relational attribute data shall be stored in a separate, well-structured relational database system. Map features and database records shall share a common unique identifier or primary key that relates a map feature to a table record.

Primary Key

The Natural Resource Database Template is an MS Access database that contains the core table "tblLocations" that contains the primary key field "LocationID". GIS data are required to also contain a field "LocationID" and be formatted the same as in the Database Template (type = Character, length = 255). The values of LocationID must be unique and less than 255 characters in length and should be consistent (perhaps based on sampling strategy). Actual value domains should be specified in the study plan or in consultation with the park, network, region, or program GIS/data manager(s).

For more information and the data dictionary describing the Natural Resource Database Template see <http://www1.nature.nps.gov/im/apps/template/index.htm>.

Other attribute fields may be included in the GIS feature attribute table if integral to symbology rendering. These fields will be duplicates of the MS Access data with fields formatted and values consistent with the database fields, and should be refreshed before final delivery of the dataset. Any such fields included in the GIS feature attribute table will be detailed in the descriptive document.

Attribute Accuracy

Every theme may have different attribute data requirements. In general, attribute data entry and quality control should follow good data management practices including verification of precise data entry and validation of possible domain values. All attribute accuracy assessments and corrective actions will be detailed in the descriptive document. Contractors or cooperators

should consult with the park, network, region, or program GIS/data manager(s) if guidance is needed about good data management practices.

Quality Control

Accuracy assessments of spatial and attribute data should include creation of check plots with spatial features labeled. The Descriptive Document will include the accuracy assessment method(s) performed and scale at which the data were collected. Results of tests used to verify all applicable horizontal, vertical and attribute accuracy measurements will be provided when data are delivered.

When the contractor has completed 10% of the spatial and attribute data development, the contractor must supply the data to the project manager for quality control purposes. The data must be delivered in conformance to the spatial data format requirements. Once the park and RTSC have checked the data and found it acceptable, the contractor may continue data development. Once the contractor has completed the work, the project manager must determine that the spatial data, attribute data, and Federal Geographic Data Committee (FGDC) compliant metadata are acceptable before the job is considered complete.

For each map feature, the estimated horizontal error in meters shall be recorded in the database field "EstHError." Estimated horizontal error is a calculation of the error range associated with a location. The required Federal reporting standard is the radius of a circle of uncertainty, such that the true or theoretical location of the point falls within that circle 95% of the time. For digitized features, the horizontal error is determined by the source map scale and digitizing precision. For GPS locations, most units provide an estimate of positional accuracy that can be used to estimate the horizontal error. An optional field describing the data source for each feature may also be included.

Metadata

All spatial data submitted shall include metadata that meets the minimum content standard for digital geo-spatial metadata (FGDC metadata; see <http://geology.usgs.gov/tools/metadata/>). (Project managers should request metadata exceeding the minimum requirements, e.g., the Biological Profile whenever appropriate.) The metadata must be parsed with no errors prior to submission using the Metadata Parser (MP) provided by the FDGC. The metadata should be delivered in FGDC-standard formatted ASCII text with a .TXT extension, hypertext markup language with an .HTML extension, and standard general markup language with an .SGML extension (<http://geology.usgs.gov/tools/metadata/>).

Specifications for the attributes and database tables attached or linked to the spatial data must be documented in the "Attribute Entity" section of the FGDC metadata and include:

- Field name
- Field description
- Field format
- Valid values

The Descriptive Document should also include a more easily readable, tabular-formatted data dictionary with the attribute and database tables specifications. The data dictionary should be listed by table and include the field name, field format, field width, and field description with valid values. An entity and relationship diagram should be included for relational tables if applicable.

Several example FGDC-compliant metadata records and browse graphics may be reviewed at http://www.nps.gov/gis/sample_data.htm for reference. To learn more about getting started with

FGDC metadata or using the MetaParser program see <http://geology.usgs.gov/tools/metadata/> or contact your project or data manager. For complete information on FGDC Metadata see www.fgdc.gov.

Note: Experienced metadata developers should estimate 4 hours per theme layer for metadata development. If the developer is unfamiliar with FGDC metadata development and guidelines, estimate an additional 20 hours for learning the process.

Legend

If project deliverables include thematic map displays, the corresponding symbology shall be included as an ArcView 3.x legend file (.AVL). Additionally, fields integral to symbolization must be present in the delivered GIS feature attributes. The descriptive document shall include a description of the thematic display and the fields required for symbol rendering.

Linked Documents

Project documents such as user manuals and detailed descriptions can be linked to map features through "hot linking". Hot linking (hyperlinking) allows the user to click a map feature and have a related document open and jump to the chapter associated with an attribute of that map feature. If an associated document is included with the intention of hot linking (hyperlinking) the following is required:

MS Word Documents (for conversion to Windows Help Files)

- The document(s) shall be an MS Word formatted file.
- The document(s) will include a table of contents with separate listings for each "topic" or description that relates to a GIS feature (e.g., extensive textual descriptions of each and every feature of a theme).
- Include a separate tabular list of which "topics" correspond to each linking field value in the GIS theme (i.e. the key values for linking the document to the GIS).

HTML Documents

- The document(s) shall be an HTML formatted file.
- The document(s) will include a table of contents with separate listings and anchors for each "topic" or description that relates to a GIS feature.
- Include a separate tabular list of which "topics" correspond to each linking field value in the GIS theme (i.e. the key values for linking the document to the GIS).

For a more on linking documents to features, see the NPS Theme Manager helpfile "Advanced Theme List Parameters" topic at <http://www.nature.nps.gov/im/apps/thmmgr/index.htm>.

Linked Graphics or Digital Photographs

If any linked digital photographs are included with the data set, they should be in a format that is directly readable in ESRI ArcView 3.x. Image types that can be directly hot linked (hyperlinked) to a theme in ArcView's Avenue scripting language include .GIF, .JPEG/.JPG, MacPaint, Microsoft DIB, Sun Raster files, .TIFF, .TIFF/LZW compressed, X-Bitmap, and .XWD. Web browsers and the NPS Graphics Viewer also allow the use of linked .GIF, .JPEG/.JPG, and .BMP formats (see <http://www.nature.nps.gov/im/apps/thmmgr/index.htm>), including large graphic images that are problematic in ArcView 3.x.

Images and graphics shall be organized in a file folder or directory structure that provides a logical hierarchical format. The directory structure recommended by the national I&M Program may be downloaded at <http://www.nature.nps.gov/im/gis/standards.htm>.

Map features with linked graphics/photographs should contain an attribute that records the absolute directory path and filename (multiple images should be separated by commas). The suggested field name is "Images." Map layers should have meaningful names that relate to the map theme and its attributes, and digital image filenames should be encoded with this value. Any file coding schemes that are used should be documented and included in the Descriptive Document.